生产过程中人的极限工作时间研究

崔铁军[1,2]，李莎莎[1]

(1. 安全科学与工程学院，辽宁工程技术大学，辽宁 阜新 123000; 2 大连交通大学 辽宁省隧道与地下结构工程技术研究中心，辽宁 大连 116028)

摘 要：操作工人在长时间的工作过程中，由于精神和身体的疲劳会产生失误，人的可靠性降低，导致生产残次品，甚至危及人身安全。为了了解在保证可靠性的前提下，操作者最长可持续的工作时间(极限工作时间)，提出了一种结合统计过程控制图(statistical process control chart，SPCC)和差分自回归移动平均模型(autoregressive integrated moving average model，ARIMA)的方法来进行确定。用连续时间中单位时间间隔内的次品率衡量人的可靠度，并形成次品率时序序列(defective rate time series，DRTS)。一方面使用该数列基于 ARIMA 构建 SPCC，并确定样本的上控制线(up control line，UCL)、下控制线(low control line，LCL)和极限时间范围；另一方面使用 ARIMA 和支持向量机(support vector machine，SVM)配合实时跟踪算法(realtime tracing algorithm，RTA)对需要确定极限工作时间(limit working time，LWT)的某一操作者的 DRTS 进行预测。最后将预测曲线置于 SPCC 中来确定该操作者 LWT 的范围。通过对 10 个操作者 8 小时，每隔 10 分钟采样得到的 DRTS 分析，表明该类工作参数的范围为 UCL=[3.89，4.66]、LCL=[-4.98，-3.90]、LWT=[393，450] min。确定了一个新操作者的 LWT 为[435，452] min，即 435 min 时该操作者就应该停止工作进行休息。

关键词：人的可靠性；差分自回归移动平均模型；统计过程控制图；极限工作时间

中图分类号：X913.4 doi: 10.3969/j.issn.1001-3695.2017.12.0852

Study on determine limit working time of human in production process

Cui Tiejun[1, 2], Li Shasha[1]

(1 *College of Safety Science & Engineering Liaoning Technical University*, *Fuxin Liaoning* 123000 *China*; 2 *Tunnel & Underground Structure Engineering Center of Liaoning*, *Dalian Jiaotong University*, *Dalian Liaoning* 116028 *China*)

Abstract: In the process of long work hours, operators can produce errors due to mental and physical exhaustion, human reliability is reduced, and it leads to production of incomplete and even endangers humanal safety. In order to understand operator's longest sustainable working time on the premise of given reliability, a method of combining SPCC and ARIMA is put forward. With unit time interval defect rate in the continuous time to measure the reliability of human, and form a defective rate time series. On the one hand, the SPCC is built using DRTS based on ARIMA to determine the sample of UCL, LCL and LWT range. On the other hand, the DRTS of an operator is predicted using ARIMA and SVM with RTA. Finally the prediction curve is put into the SPCC to determine the LWT range. Based on 8 hours, 10 operators DRTS analysis of sampling for every 10 minutes, indicates that the scope of this kind of working parameters for the UCL = [3.89, 4.66], LCL = [4.98, 3.90], LWT = [393450]. Identified a new operator LWT is [435452] minutes, 435 minutes when the operator should stop working to have a rest.

Key words: reliability of the human; ARIMA; SPCC; limit Working time

0 引言

我国是一个劳动力大国，在高度机械化和自动化的今天，仍有大量的劳动者从事着以人为主要动力来源的劳动。人与机器相比有着丰富的创造能力，但同时人在长时间的劳动过程中会在生理和心里方面产生疲劳。在疲劳状态下工作的操作人员，心情烦躁，注意力难以集中，不但工作效率低下，而且会生产较多残次品。这样会加重操作人员的心里压力，使人的行为更加不可靠，甚至会由于操作不当造成人身伤亡。所以，确定一个适当的工作时间，不仅能提高操作者的效率，而且能提高产品的合格率。一方面保护了劳动者，另一方面也增加了企业的经济效益。即在保证操作人员操作可靠性不小于某个特定值的

收稿日期：2017-12-14；**修回日期：**2018-01-25

作者简介：崔铁军（1983-），男，辽宁沈阳人，副教授，博士，主要研究方向为系统可靠性及力学系统稳定性（ctj.159@163.com）；李莎莎（1988-），女，辽宁盘锦人，博士研究生，主要研究方向为安全系统工程、系统可靠性.

前提下，确定操作者可以工作的最长时间，即极限工作时间(LWT)。

关于人的可靠性及适合的工作时间的研究，目前有柴松等人[1,2]的基于 CREAM 和不确定推理的人因可靠性分析方法和海洋工程人因可靠性定量分析方法与应用；董学军等[3]的基于补偿和不可替代因素合成的人因可靠性分析方法；蒋建军等人[4,5]的基于隐马尔可夫的核电厂半数字化人-机界面事故诊断过程人因可靠性模型和考虑人因的核电厂主控室认知可靠性模型研究；倪蓉[6]的煤矿职工疲劳度相关因素分析与控制研究。

笔者参考了文献[7~9]，并根据自己的经验[10~14]，提出了基于 SPCC 和 ARIMA 的人的 LWT 的确定方法。该方法主要分为两部分，一部分是构造 SPCC 并得到 LCL、UCL、中心线(center line，CL)和极限时间范围，这部分的数据来源是多个样本（操作者）的统计数据；另一部分使用 ARIMA，配合 SVM 和 RTA 预测某一个新样本在工作时间不足的情况下后继时间序列的变化，进而形成该新样本的 DRTS，结合得到的 SPCC 可预测该样本的 LWT。

1 方法论

该方法涉及到的相关知识有 SPCC、ARIMA，SVM 和 RTA。算法作用有两个，一是统计多个样本（操作者）确定某种操作的泛化 LCL、UCL、CL 和 LWT；二是基于这些统计结果预测某个个体（操作者）的 LWT。SPCC 对多个样本统计数据进行整理，得到数据的 LCL、UCL、CL 和 LWT。可得数据变化趋势是否将超限，从而可根据统计数据的特征，预测同类型数据是否有超限的可能，形成 SPCC 图是衡量标尺。ARIMA 是算法核心，SPCC 处理时间序列。ARIMA 在形成 SPCC 时适用于无相关性的时间序列数据。为了预测新操作者的 LWT，需要对短时间内的新操作者的 DRTS 进行分析，并对未来的序列数据进行预测。但 DRTS 很可能是线性和非线性的序列数据叠加的，根据文献[9]可将线性数据使用 ARIMA 进行预测，非线性数据使用 SVM 预测。根据文献[8]可使用 RTA 对时序数据进行跟踪，以保证预测的精确性。经过上述处理的线性和非线性时序数据相叠加形成最终的该操作者的 DRTS 数据，放入 SPCC 图中，即可确定他的 LWT。算法流程如图 1 所示。

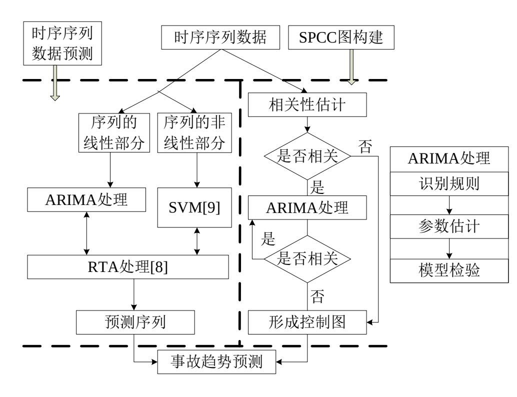


图 1 算法的流程图

从图 1 可知算法分为左右两部分用粗虚线分开。左边为时序数据的预测部分，右面为 SPCC 图构造部分。最右侧的 ARIMA 处理模块包括了三个部分，左右两部分都要用到所以独立开来。右侧形成 SPCC 图，是来源于多个样本的统计结果，是整个算法的基础。SPCC 可得到 LCL、UCL、CL 和 LWT，进而处理单个样本的 LWT 预测。左侧为预测单个样本 DRTS 数据，在 SPCC 中与 LCL 和 UCL 比较，最终确定 LWT。

2 方法步骤及其解释

SPCC 对实时数据进行监控和预测。SPCC 基于 ARIMA 与传统的 SPCC 是不同的[15,16]。SPCC 构造过程分三步：相关性评价；ARIMA 构建；控制图构建。自相关函数(autocorrelation coefficient function，ACF)在第一步中，用来评价时间序列的自相关性。如果不相关直接执行第三步，用原始时间序列构建控制图。否则用 ARIMA 消除相关性，在第二步产生时间序列的非相关残差。第三步用时间序列残差构建控制图。

2.1 相关性估计

使用 ACF 估计相关性。设滞后时间间隔为 k，使用 MATLAB 中的 autocorr 计算相关性。如果 ACF 图显示的数值绝对值总体趋势减小，则断定时间序列是相关[17]。

2.2 ARIMA 构建

如果数据是相关的，那么使用 ARIMA 消除来消除相关性。ARIMA 模型可表示为 ARIMA(p,d,q)，其中 p,d,q 为大于等于 0 的整数，p 为自回归项；d 为时间序列成为平稳时所做的差分次数；q 为移动平均项数。给定 ARIMA(p,d,q)，原时间序列 x_t 可被拟合并转换为一个白噪声过程 e_t，之间的关系如式（1）所示[7]。

$$\left(1-\sum_{p=1}^{p}\varphi_p B^p\right)(1-B)^d x_t=(1+\sum_{q=1}^{q}\varphi_q B^q)e_t \tag{1}$$

其中：φ_p 和 φ_q 自回归系数和移动平均系数，B 延迟算子，$B^j x_t = x_{t-j}$，j 为延迟次数。

ARIMA 的构造分三个步骤，即识别规则，参数估计和模型检验[18]。

a)在识别规则中，使用逐次差分将时间序列转换为稳定数据，一次和二次差分公式分别如式(2)(3)所示[7]。

$$(1-B)x_t = x_t - x_{t-1} \tag{2}$$

$$(1-B)^2 x_t=(1-2B+B^2)x_t = x_t-2x_{t-1}+x_{t-2} \tag{3}$$

ACF 和偏自相关函数(partial autocorrelation coefficient function，PACF)可绘制出各种差分时间序列的对应曲线。如果其他参数不变，差分次数 d 的变化对 ACF 和 PACF 形成的图形影响较小[19]。具体来说，ACF 图像拖尾和 PACF 延迟 p 后截尾时应使用 ARIMA，具体方法见文献[19]。

b)使用最大化条件对数似然函数 $l_*(\varphi_p,\theta_q,\delta_{e_t}^2)$ 进行估计，如式(4)所示[7]。

$$l_*(\varphi_p,\theta_q,\delta_{e_t}^2)=-\frac{n}{2}\ln(\delta_{e_t}^2)-\frac{S_*(\varphi_p,\theta_q)}{2\delta_{e_t}^2} \tag{4}$$

其中：$\delta_{e_t}^2$ 是 e_t 的方差；$S_*(\varphi_p,\theta_q)$ 通过式(5)进行估计。

$$S_*(\varphi_p,\theta_q)=\sum_{T=1}^{n}e_t^2(\varphi_p,\theta_q\,|\,(1-B)x_{t_*},e_{t_*},(1-B)x_t\,) \tag{5}$$

如果，ARIMA(p,d,q)中 $p\neq 0,q\neq 0$，通过最小艾卡信息量准则(lowest Akaike information criterion，AIC)进行定阶[20]，其具体表达式如式(6)所示。

$$AIC(t,j)=\ln(\delta_{e_{t,j}}^2)+\frac{2(t+j)}{N} \tag{6}$$

其中：$\delta_{e_{t,j}}^2$ 是 e_t 的方差的估计值，N 为样本数。

c)检验时间序列的相关性是否消除。使用 ACF 对白噪声 e_t 进行分析可进行验证。当延迟 j 大于 10，且 ACF 的分析结果趋近于 0 时可认为相关性已被消除[21]。否则重复进行 ARIMA 处理，直到满足上述条件。

d)利用 RTA 改进 ARIMA[8]。在 ARIMA 基础上提出与实时跟踪算法 RTA 结合的改进模型。假设要对时间序列 $\{x_i\,|\,i=1:t\}$ 进行预测，假设步长为 m，利用 $\{x_i\,|\,i=1:t\}$ 预测 $\{x_i\,|\,i=t+1:t+m\}$，去掉 $\{x_i\,|\,i=1:t+m\}$ 的前 m 个数据，形成 $\{x_i\,|\,i=m+1:t+m\}$ 进行下一步的时序数列预测。

e) ARIMA 和 SVM 模型分别在处理线性和非线性数据结构的预测问题时有着各自的优势。可将 ARIMA 和 SVM 结合使用，以求达到增强预测模型对数据样本的适应能力，从而提高预测的效果[9]。

2.3 控制图构建

构建休哈特统计控制过程图(Shewhart statistical process control chart，SSPCC)，基于 Shewhart three sigma control theory[22] 来确定中心线 CL。历史数据可为无相关性的原时间序列 x_t 或 ARIMA 处理后的白噪声过程 e_t。α_i 为在时间点 i 处的历史数据。控制上线 UCL、控制下线 LCL 和中心线 CL 如式（7）所示。

$$UCL/LCL=2CL=\bar{X}\pm 3\sigma=\bar{X}\pm 3\frac{\overline{mr}}{d_n} \tag{7}$$

其中：$\bar{X}=\frac{\sum_{i=1}^{T}\sigma_i}{T}$，T 为时间序列长度的平均值。$\overline{mr}$ 表示平均移动范围，$\overline{mr}=\frac{\sum_{i=2}^{T}mr_i}{T-1}$，其中 $mr_i=\alpha_i-\alpha_{i-1}$。

SSPCC 使用时间序列中两个连续值作为移动范围[22]，所以 n=2, $d_2=1.128$ [23]。

3 实例计算

在流水线上的操作人员往往由于工作时间长而产生疲劳，导致生产的部件残次品增多，甚至可能由于操作不当对自身安全遭威胁。监测 10 个工人在 8 小时内的生产残次品比例（次品率）作为衡量操作人员操作可靠性指标。取样时间间隔为 10 分钟，次品率为 10 分钟内生产的残次零件/全部零件。构成了有 48 个数据点的 10 个样本时序序列数据。根据次品率统计，在 8 小时内的 10 个工人的次品率在 0~0.2，曲线是起伏的，但总趋势递增。所以采用一阶差分对时序序列进行处理。由于一阶差分、ACF、PACF 在 MATLAB 中已有工具包，这里不在说明。给出 10 个样本中第一个样本的数据处理过程的图示，如图 2 所示。

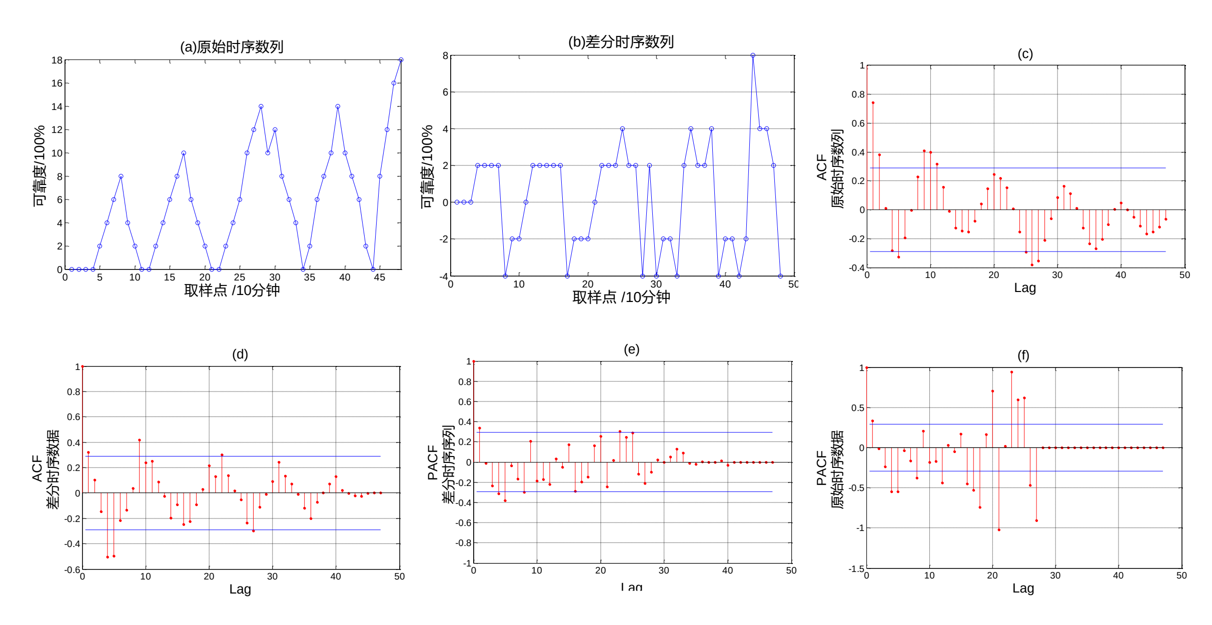


图 2　第一个样本的数据处理过程图

图 2 中，(a)表示了该操作者（第一个样本）的 DRTS 数据，一共 48 个采样点，480 分钟的工作时间。根据次品率定义，可知 DRTS 数据是起伏波动的，随时间增加总体增加的趋势。由图(c)和(f)可见，大部分数据的 ACF 和 PACF 值都超出允许范围，时间序列的平稳性不满足计算要求，需要进行差分计算。通过差分计算的一次差分序列、ACF 和 PACF 分别如图(b)(d)和

(e)可见一次差分序列值趋于稳定，ARIMA 模型中的系数 d 定为 1。然后利用 AIC 来确定 ARIMA 模型的阶数 p 和 q。通过计算发现当 p=4,q=4 时 AIC 值最小，故使用模型 ARIMA(4,1,4)。第一个操作工人的差分 DRTS 的 SPCC 如图 3 所示。

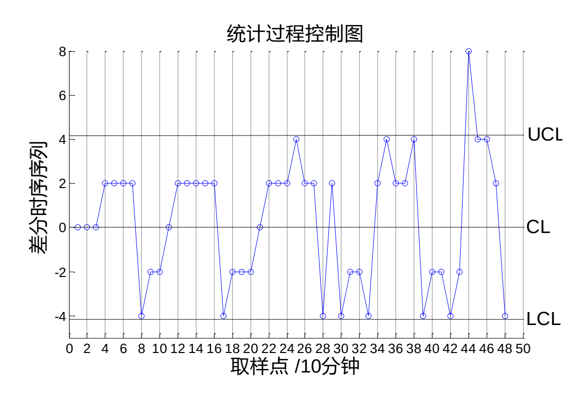


图 3　第一个操作工人的差分次品率时序序列的 SPCC

从图 3 可知第一个操作工人的 CL=0、UCL=4.17、LCL=-4.17。差分 DRTS 曲线与 UCL 交于采样点 43~44，即在 430~440 min 该操作人员达到限值，不适宜继续工作。则 LWT 范围就是 430~440 min（LWL=437 min）。同理确定 10 个操作人员的 LWT 如表 1 所示。将表 1 中参数的值进行综合得到了该类工作的对应参数的范围，即 UCL=[3.89,4.66]、LCL=[-4.98,-3.90]、LWT=[393,450] min。

表 1　10 个操作人员的极限工作时间

指标	1	2	3	4	5	6	7	8	9	10	范围
UCL	4.17	4.08	3.89	5.11	4.66	4.54	3.90	4.43	3.68	3.99	[3.89,4.66]
LCL	-4.17	-4.08	-4.03	-4.98	-4.66	-4.34	-3.90	-4.33	-4.09	-4.18	[-4.98,-3.90]
CL	0	0	-0.07	0.065	0	0.1	0	0.05	-0.205	-0.095	
LWT/分	437	404	401	450	422	430	393	425	397	406	[393,450]

注:范围确定是取 10 个样本对应值的最大和最小值,作为区间的两个端点。

上述过程完成了图 1 中右侧框图部分，即完成了 SPCC 的构建。下面是判断某一新样本（新操作者）的 LWT。

对于一个新操作者（不在样本的 10 个操作者内），利用以进行工作时间的 DRTS 的一次差分序列对该操作者进行 LWT 预测。利用文献[8,9]的方法进行时序序列的拟合和预测，得到了较好的预测结果。主要思想是将原始时序序列拆分为线性序列和非线性序列，线性序列由 ARIMA 处理并进行预测，处理过程与上述过程相近，见参考文献[8]；非线性序列可由 SVM 处理并进行预测，见参考文献[9]。同时在处理过程中结合 RTA 方法进行数据跟踪，提高预测精度。最后将上述线性和非线性序列进行叠加，达到对原 DRTS 数据预测的效果。实际值与预测值的残差如图 4 所示。预测误差为 2.459%。该操作者的实际 DRTS 数据曲线和预测曲线如图 5 所示。

在图 5 中，底纹区域 UCL 和 LCL 是通过原始 10 个样本确定的。使用预测方法对该操作者的 DRTS 预测较为成功（误差为 2.459%）。而且使用预测曲线结合以得到的 UCL 和 LCL 区域可确定该操作者的 LWT。如图中的竖向底纹区域，该区域的水平方向区间宽度是由 DRTS 的预测曲线与 UCL 相交的区域宽度确定的。对于该操作者的 LWT 为[435,452]分钟，即 435 分钟时该操作者就应该停止工作进行休息。

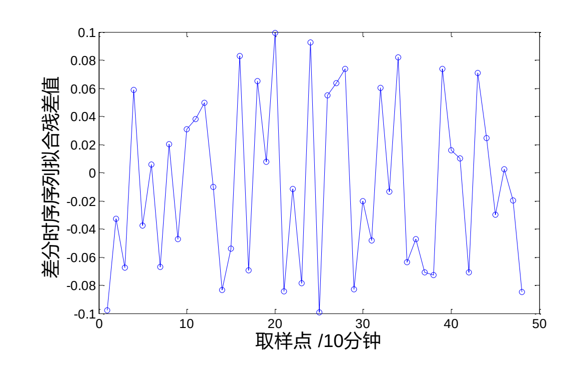


图 4　次品率时序序列数据差分拟合的残差

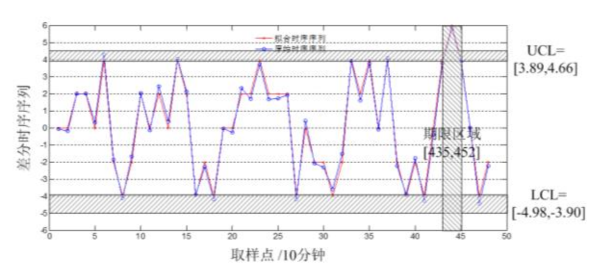


图 5　该操作者的 LWT 分析图

4　结束语

a）提出了极限工作时间确定方法。分为两部分，一是构造 SPCC 并得到 LCL、UCL、CL 和极限时间范围，数据来源是多个样本（操作者）的统计；二是使用 ARIMA，配合 SVM 和 RTA 预测某一个新样本在工作时间不足的情况下后继时间序列的变化，进而形成该新样本的 DRTS。结合得到的 SPCC 可预测该样本的 LWT。

b）基于 10 个样本形成了 SPCC。使用模型 ARIMA(4,1,4) 进行 SPCC 构建。得到了 10 个样本（操作者）的 UCL、LCL、CL 和 LWT。综合得到了该类工作的对应参数范围，即 UCL=[3.89,4.66]、LCL=[-4.98,-3.90]、LWT=[393,450] min。

c）确定了新样本（新操作者）的极限工作时间。预测误差为 2.459%，确定了该操作者的 LWT 为[435,452] min，即 435 min 时该操作者就应该停止工作进行休息。

参考文献：

[1] 柴松，余建星，马维林，等. 基于 CREAM 和不确定推理的人因可靠性分析方法 [J]. 天津大学学报, 2012, 45 (11): 958-962.

[2] 柴松，余建星，杜尊峰，等. 海洋工程人因可靠性定量分析方法与应用 [J]. 天津大学学报, 2011, 44 (10): 914-919.

[3] 董学军，陈英武. 基于补偿和不可替代因素合成的人因可靠性分析方

法 [J]. 系统工程理论与实践, 2012, 32 (9): 2087-2096.

[4] 蒋建军, 张 力, 王以群, 等. 基于隐马尔可夫的核电厂半数字化人-机界面事故诊断过程人因可靠性模型 [J]. 核动力工程, 2012, 33 (5): 79-82.

[5] 蒋建军, 张力, 王以群, 等. 考虑人因的核电厂主控室认知可靠性模型研究 [J]. 核动力工程, 2012, 33 (1): 66-72.

[6] 倪蓉. 煤矿职工疲劳度相关因素分析与控制研究 [J]. 煤矿安全, 2013, 44 (10): 230-233.

[7] Hu Xinyao, Qu Xingda. An individual-specific fall detection model based on the statistical process control chart [J]. Safety Science 2014, 64: 13-21.

[8] 秦摇鹏, 苏怀智, 沈跃军. 基于 ARIMA-RTA 组合模型的海堤工程沉降预测 [J]. 水利水运工程学报, 2013, 5: 66-70.

[9] 黄彬彬, 谷波, 任能. 基于 ARIMA-SVM 模型的翅片管蒸发器结霜性能预测 [J]. 上海交通大学学报, 2009, 43 (10): 1622-1631.

[10] 崔铁军, 马云东. 基于泛函网络的周期来压预测方法研究 [J]. 计算机科学, 2013, 40 (6A): 242-246.

[11] Cui Tiejun, Li Shasha. Study on the relationship between system reliability and influencing factors under big data and multi-factors [J]. Cluster Computing-The Journal of Networks Software Tools and Applications, https://doi. org/10. 1007/s10586-017-1278-5.

[12] Li Shasha, Cui Tiejun, Li Xingsen, *et al.* Construction of cloud space fault tree and its application of fault data uncertainty analysis [C]// Proc of International Conference on Machine Learning and Cybernetics. 2017: 195-201.

[13] Li Shasha, Cui Tiejun, Liu Jian. Study on the construction and application of cloudization space fault tree [J]. Cluster Computing: The Journal of Networks Software Tools and Applications, https://doi. org/10. 1007/s10586-017-1398-y

[14] 崔铁军, 马云东. 多维空间故障树构建及应用研究 [J]. 中国安全科学学报, 2013, 23 (4): 32-37.

[15] Alwan L C, Roberts H V. Time-series modeling for statistical process control [J]. Journal of Business and Economic Statistics, 1988, 6 (1): 87–95.

[16] Haridy S, Wu Z. Univariate and multivariate control charts for monitoring dynamic-behavior processes: a case study [J]. Journal of Industrial Engineering and Management, 2009, 2 (3): 464–498.

[17] Mohammed M A, Worthington P, Woodall W H. Plotting basic control charts: tutorial notes for healthcare practitioners [J]. Quality and Safety in HealthCare, 2009, 17 (2): 137-145.

[18] Box G E P, Jenkins G M. 1976. Time series analysis: forecasting and control [M]. 5th ed. Hoboken: Wiley, 2015.

[19] Montgomery D C, Jennings C L, Kulahci M. Introduction to time seriesanalysis and forecasting [M]. 2nd ed. Hoboken: Wiley-Interscience, 2015.

[20] 何林, 欧进萍. 基于 ARMAX 模型及 MA 参数修正的框架结构动态参数识别 [J]. 振动工程学报, 2002, 15 (1): 47-51.

[21] Nau R. Introduction to ARIMA: non-seasonal models [EB/OL]. (2014-11-29). https://people.duke.edu/~rnau/Slides_on_ARIMA_models-Robert_Nau.pdf.

[22] Shewhart, W. A. , 1931. Economic Control of Quality of Manufactured Product [M]. D. VanNostrand Company Inc. , New York.

[23] Montgomery D C. Introduction to statistical quality control [M]. 6th ed. Hoboken: Wiley, 2009.